



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

Address: COMMISSIONER FOR PATENTS

P.O. Box 1450

Alexandria, Virginia 22313-1450

www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/670,245	09/26/2003	Masaru Sugano	031198	8594

38834 7590 02/16/2011  
WESTERMAN, HATTORI, DANIELS & ADRIAN, LLP  
1250 CONNECTICUT AVENUE, NW  
SUITE 700  
WASHINGTON, DC 20036

EXAMINER

PRINCE, JESSICA MARIE

ART UNIT

PAPER NUMBER

2482

NOTIFICATION DATE

DELIVERY MODE

02/16/2011

ELECTRONIC

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentmail@whda.com

### Office Action Summary

**Application No.**

10/670,245

**Applicant(s)**

SUGANO ET AL.

**Examiner**

JESSICA PRINCE

**Art Unit**

2482

**Period for Reply** -- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 06 December 2010.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-23 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-23 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-846)
- 3) ☐ Information Disclosure Statement(s) (PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: \_\_\_\_\_

## **DETAILED ACTION**

### ***Response to Arguments***

1. Applicant's arguments filed 12-06-2010 have been fully considered but they are not persuasive.

As to Applicants argument that Chakraborty is not able to detect a commercial scene which includes a plurality of cuts from the video because Chakraborty is not able to classify each shot into a commercial scene, that is Chakraborty is not able to recognize whether each detected shot is part of the commercial scene.

The examiner respectfully disagrees. Chakraborty discloses video are playing an increasingly import role in education and commerce, Column 1 line 15-18. Further, when the approximate maximum duration is known, since the frames/sec is always known, the maximum frame duration for the scene change is readily ascertainable. If any of the windows have a duration that exceeds this threshold, it may be assumed that the window in question is not likely to be a gradual scene change. In such as case, further examination becomes necessary. The possibilities are that either the window represents just motion or a combination of scene change and motion. In the preferred embodiment, if any window has a duration that exceeds the predefined threshold, it is assumed that the window represents motion, and consequently all points in such window are turned "off" (step 224). All the remaining windows are then identified as candidates for gradual scene change, column 14 line 20-35. Chakraborty teaches a predefined shot duration (column 13 line 15 to 35); which is equivalent to the shot

density. Therefore, since Chakraborty discloses videos in education and commerce, and based on the predefined window threshold, the scene is either gradual or abrupt, it is clear to the examiner that Chakraborty is fully capable of detecting a commercial scene based on the shot density, which reads upon the claimed limitation). In addition, Chakraborty teaches where a "shot" or "take" in video parlance refers to a contiguous recording or one or more video frames depicting a continuous action in time and space. Typically, transitions between shots (also referred to as "scene changes" or "cuts") are created intentionally by film directors, see col. 1 line 35-44. The examiner notes that a scene is a plurality of shots. Since a shot refers to a continuous recording of one or more video frames depicting a continuous action in time and space, and a scene is a plurality of shots, clearly, a scene is fully capable of including a plurality of shots containing multiple transitions (cuts) between the shots. In addition, in response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., "able to classify each shot into a commercial scene") are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

As to Applicants argument that Chakraborty is not able to detect a target scene because from the video because Chakraborty lacks the second step. Chakraborty, even

a sequence of plurality of shots which has same features because the shots constitute identical scene is not recognized as identical.

In response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., "detecting a target scene" and "plurality of shots which has same features because the shots constitute identical scene is not recognized as identical") are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

As to Applicants argument that a scene in Chakraborty is not portion of video that is defined differently from the shot, and certainly nowhere does Chakraborty define a classifying a shot into a scene composed of a plurality of shots.

The examiner respectfully disagrees. Chakraborty teaches where a "shot" or "take" in video parlance refers to a contiguous recording or one or more video frames depicting a continuous action in time and space. Typically, transitions between shots (also referred to as "scene changes" or "cuts") are created intentionally by film directors, see col. 1 line 35-44. The examiner notes that a scene is a plurality of shots. Since a shot refers to a continuous recording of one or more video frames depicting a continuous action in time and space where the transitions between shots are called cuts and a scene is a plurality of shots, clearly, a scene is fully capable of including a plurality of shots containing multiple transitions (cuts) between the shots.

As to Applicants argument that Chakraborty is silent regarding classifying the shots in the shot list into specific types of scenes.

In response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., "classifying the shots in the shot list into specific types of scenes") are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

As to Applicants argument that none of the cited references disclose or suggest any of the features recited in claims 1, 4, 9, 13, and 14, other than the claimed 'a shot segmentation device to segment the video into respective shots', because none of the reference teach or suggest performing the claimed operations on the segmented shots (i.e., on shots after they are segmented), and none of the references teaches performing the claimed operations on segmented shots to classify the shot into a specific type of scene including a plurality of continuous shots.

The examiner respectfully disagrees. The Examiner respectfully disagrees. It is the combination of the Chakraborty and Toklu as a whole that teaches the claimed limitations. Further, Chakraborty (modified by Toklu) as a whole disclose to perform the limitations of the claims on segmented shots and classifying a scene including a plurality of continuous shots. Toklu discloses where conventional video summarization

method typically include segmenting a video into an appropriate set of segments such as video "shots", Col. 1 line Further, taught is a video segmentation module 12 partitions a video file (that is either retrieved from the database 11 or input in real-time as a video data stream) into a plurality of video segments (or video segments) and then outputs segment boundary data corresponding to the input video data. It is to be understood that that any conventional process may be employed herein for partitioning video data which is suitable for implementation with the present invention. The above-incorporated cut detection method partitions video data into a set of "shots" comprising visually abrupt cuts or camera breaks. As stated above, a video "shot" represents a contiguous recording of one or more video frames depicting a continuous action in time and space, col. 5 line 39-60. Since Toklu discloses to the method partitions video data into a set of "shot" comprising visually abrupt cuts or camera break, it is clear to the Examiner that Toklu segments or classifies continuous shots as either abrupt or as a camera break, which reads upon performing operations segmented shots and classifying continuous shots. Therefore, Chakraborty (modified by Toklu) discloses to perform segmentation of video into shots, classify the shots as well as classify the shots before performing the operations of Chakraborty. In response to applicant's arguments against the references individually, one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.*, 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986).

As to Applicants argument that Yilmaz does not use a number of shot boundaries in a predetermined interval to classify a commercial scene. Unlike the claimed invention, Yilmaz uses a means of eigenvectors in a shot to label the shot as an advertisement.

The examiner respectfully disagrees. Chakraborty discloses to detect the commercial scene (abrupt scene, and video in commerce contains commercial scenes) by scene change (shot boundary) by comparing each of the computed metrics for the successive frames to threshold levels associated with the respective difference metrics, col. 5 line 20-24. Further disclosed is that this threshold level is user defined because such threshold depends on the type of film being processed. When the approximate maximum duration is known, since the frames/ sec is always known, the maximum frame duration for the scene change is readily ascertainable. If any of the windows have a duration that exceeds this threshold, it may be assumed that the window in question is not likely to be a gradual scene change see col. 14 line 19-27. Therefore, it is clear to the examiner that Chakraborty discloses to determine abrupt (commerce scenes) using the scene change (shot boundary) where the boundary is compared to for the successive frames to a threshold that has a user defined duration (predetermined interval). Yilmaz discloses to cluster news video into news and advertisement based on the shots boundaries. Therefore, substituting the explicit teaching of Yilmaz to detect a commercial scene based on shot boundaries with Chakraborty, now discloses the claimed limitation. Thus, the combination of Chakraborty modified by Yilmaz discloses the claimed feature. In response to applicant's arguments against the references



individually, one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.*, 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986).

As to Applicants argument that Gonsalves does not disclose or suggest anything about an "inserting means [that] makes a type of the video transition effect to be inserted different according to whether the highlight scenes to be combined are the dynamic scene or the static scene.

The Examiner respectfully disagrees. It is the combination of Nakamura (modified by Pan and Gonsalves) that teaches applicants limitation. In this case, Gonsalves teaches allowing the video editor to insert a video transition effect on a field/frame by field/frame basis in order to improve accuracy of the effect, see col. 3 line 11-14, and 24, between two frames, col. 4 line 65-67, col. 5 line 50-52 and fig. 3b:320A-320b). Taking the teachings of Nakamura (modified by Pan) where Pan discloses where a special effect, or edit effect at block 16, is almost always present between the normal shots in block 12 and 16 and the slow motion replay segment in block 18. After the slow motion replay in block 18, another edit effect in block 20, is usually present before resuming normal play, [0020]. The edit effects in 20 and edit effects out 30, mark the starting and end points of the procedure of the slow motion replay 14, and typically are gradual transitions, such as fade in/out, cross/additive-dissolve, and wipes, [0030] with the teachings of Gonsalves where it is disclosed to implement special effects on a

Art Unit: 2482

field/frame by field/frame basis, it is clear to the Examiner that the combination is fully capable and suggest to insert special effect on a frame by frame or field by field basis, where inserted edit effects for slow motion replay (static highlight scene with little motion) are gradual. Since there is following the action shot a normal shot, and almost always there is an edit effect between the normal shot and slow motion replay segment, and edit effects in 20 and edit effects out 30, mark the starting and end points of the procedure of the slow motion replay 14, and typically are gradual transitions, such as fade in/out, cross/additive-dissolve, and wipes, it is clear to the Examiner that for the slow motion replay (static highlight scene), the effects in and out are gradual, which reads upon the claimed limitation.

As to Applicants argument that Pan is silent as to what parameters inset of the effect is based. That is Pan does not disclose an inserting means that includes " a dynamic/static scene detector to detect whether a highlight scene is a dynamic scene with much motion or a static scene with little motion. Accordingly, the combination of the teachings of Nakamura, Pan, and Gonsalves does not result in the claimed limitation.

The Examiner respectfully disagrees. Pan discloses where a special effect, or edit effect at block 16, is almost always present between the normal shots in block 12 and 16 and the slow motion replay segment in block 18. After the slow motion replay in block 18, another edit effect in block 20, is usually present before resuming normal play, [0020]. The edit effects in 20 and edit effects out 30, mark the starting and end points of

Art Unit: 2482

the procedure of the slow motion replay 14, and typically are gradual transitions, such as fade in/out, cross/additive-dissolve, and wipes, [0030]. Since there is following the action shot a normal shot, and almost always there is an edit effect between the normal shot and slow motion replay segment, and edit effects in 20 and edit effects out 30, mark the starting and end points of the procedure of the slow motion replay 14, and typically are gradual transitions, such as fade in/out, cross/additive-dissolve, and wipes, it is clear to the Examiner that for the slow motion replay (static highlight scene), the effects in and out are gradual, which reads upon the claimed limitation.

### ***Claim Rejections - 35 USC § 112***

The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

2. Claims 1, 4, 9, 1-14, 21, 23 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. The examiner is unable to find

Art Unit: 2482

support in the disclosure as originally filed for the claim limitation "wherein each scene includes a plurality of cut points".

***Claim Rejections - 35 USC § 103***

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

5. Claims 1-3, 15, and 17-20 and 23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chakraborty et al., US-7, 110,454 in view of Toklu et al., US-6,549,643 and further in view of Park et al., US-6,597,738.

Re **claim 1**, Chakraborty discloses a scene classification apparatus of video for classifying a sequence of shots into a dynamic scene with much motion or a static scene with little motion, where the dynamic scene and the static scene respectively include a plurality of continuous shots and are thus a larger unit than a shot, comprising: a calculator for calculating shot density (histogram difference metric, a histogram is a graphical display of tabulated frequencies and fig. 2A: 203) DS of the video per a time

unit(fig. 3C) the from the respective shots (extracted video frames, fig. 2); a calculator for calculating motion intensity (histogram difference metric, a histogram is a graphical display of tabulated frequencies and fig. 2A:203. Further regarding claim 2A, the element 212 output the potential shot/scene change location based on histogram difference, therefore, it is clear to the Examiner that Chakraborty teaches to disclose the density, which reads upon the claimed limitation) ; a calculator for calculating motion intensity of the respective shot (fig. 2A, element 211, outputs potential shot/scene change locations based on interframe difference. Since fig. 2A, element 211, outputs potential shot/scene change locations based on interframe difference, it is clear to the Examiner that Chakraborty discloses to calculate the motion of the shot, which reads upon the claimed limitation) of the respective shots (extracted video frames, fig. 2); and a dynamic/static scene classifier (metric computation col. 5 line 9-11, fig. 1:14-17 and fig. 2A) for classifying the sequence (continuous units or "shots" col. 1 line 35-37) of shots into the dynamic scene (abrupt scene, see abstract, furthermore, the meaning of abrupt is interpreted as sudden or fast) with much motion or the static scene with little motion (gradual scene, see abstract, furthermore, the meaning of gradual is interpreted as slow and not moving quickly) based on the shot density (histogram difference, a histogram is a graphical display of tabulated frequencies) and the motion intensity of the respective shots (Chakraborty discloses where the output of each of the scene change detection processes are potential shots/scene change location based on the respective metrics (steps 211,212, and 213), both abrupt and gradual. The next steps in the preferred scene change detection process involve identifying and validating the scene

changes based on various conditions. For instance, referring to FIG.2b, in the preferred embodiment, abrupt scene changes are identified where candidate scene change location output from the shot detection processes of the interframe and histogram difference metrics are in agreement (step 214). In particular, abrupt scene changes are identified by verifying that the conditions regarding both the interframe difference metric and the difference are satisfied. It is to be appreciated that by integrally utilizing the scene change candidates output from such shot detection processes, false alarms in identifying scene changes that may occur due to small motion where the interframe difference is high (and thus exceed the threshold in equation 11 above) will not occur since the condition for the histogram difference for the candidate must also be satisfied. (in which case, for small motion, such condition typically will not be satisfied), Column 12 line 60 to column 13 line 14 and fig. 2A-2B. Further regarding fig. 2A, element 211, outputs potential shot/scene change locations based on interframe difference, it is clear to the Examiner that Chakraborty discloses to determine if the scene is gradual or static, which reads upon the claimed limitation) wherein each scene includes a plurality of cut points (Chakraborty teaches where a "shot" or "take" in video parlance refers to a contiguous recording or one or more video frames depicting a continuous action in time and space. Typically, transitions between shots (also referred to as "scene changes" or "cuts") are created intentionally by film directors, see col. 1 line 35-44. The examiner notes that a scene is a plurality of shots. Since a shot refers to a continuous recording of one or more video frames depicting a continuous action in time and space where the transitions between shots are called cuts and a scene is a plurality of shots, clearly, a

scene is fully capable of including a plurality of shots containing multiple transitions (cuts) between the shots).

Chakraborty does not explicitly teach segmented shots; shot segmentation device to segment the video into respective shots based on a cut points. However, Toklu teaches a shot segmentation device to segmentation device to segment the video into respective shots based on a cut points (video segmentation module 12, column 5 line 38-57, col. 6 line 54-60, and fig 1 element 12).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Toklu with Chakraborty to generate a content based visual summary of video and facilitate digital video browsing and indexing, column 3 line 40-43).

Chakraborty (modified by Toklu) is silent in regards to calculating motion intensity per unit region on the image using a value of motion vectors.

However, Park teaches calculating motion intensity per unit region on an image using a value of motion vectors (Park discloses where the motion vector values MVoz, MVoy, which are generated in the motion search value divergence processing part 2, are input (S14), the motion intensity computing part 13 obtains the motion intensity Lmv by using the formula (2) wherein, p is the quantization factor of intensity. If the computed motion intensity Lmv is smaller than a threshold value THL (S34), the motion vector comparison/conversion part 23 regards that there is a motion in the image which is a too small intensity to visualize or random noise occurs in the image during image obtaining or processing, so that the motion vector comparison/conversion part 23

converts all estimation motion values  $MV_{fx}$  and  $MV_{fy}$  into 0 (S44). On the other hand, if the computed motion intensity  $L_{mv}$  is larger than the threshold value  $THL$ , the motion vector comparison/conversion part 23 converts the motion vector values  $Mv_x$ ,  $Mv_y$  into all estimated motion values  $MV_{fx}$  and  $MV_{fy}$  (S54), col. 16 line 20-49 and fig. 14).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Park with Chakraborty (modified by Toklu) for to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have composite information attribute, thereby being suitable for efficient data search of multimedia (col. 1 line 30-36).

Regarding **claim 2**, Chakraborty (modified by Toklu and Park) as a whole teaches everything as claimed above, see claim 1. In addition Chakraborty discloses the scene classification apparatus of video according to claim 1, wherein the dynamic/static (metric computation col. 5 lines 9-11, fig. 1:14-17 and fig. 2A) scene classifier classifies a sequence of shots whose shot density (histogram difference, a histogram is a graphical display of tabulated frequencies) is larger than first reference density and whose motion intensity is stronger than first reference intensity (frame to frame intensity col. 1 lines 50-53) into the dynamic (abrupt col. 12, line 67; col. 13 line 1-3) scene.

Regarding **claim 3**, Chakraborty (modified by Toklu and Park) as a whole teaches everything as claimed above, see claim 1. In addition Chakraborty discloses



the scene classification apparatus of video according to claim 1, wherein the dynamic/static scene detector (metric computation col. 5 lines 9-11, fig. 1:14-17 and fig. 2A) classifies a shot whose shot density (histogram difference, a histogram is a graphical display of tabulated frequencies) is smaller than second reference density and whose motion intensity (histogram difference computation fig. 1:16) is weaker than second reference intensity into the dynamic scene (gradual scene).

(2) Regarding **claim 15**, Chakraborty (modified by Toklu and Park) as a whole teaches everything as claimed above, see claims 1 or 4. In addition, Chakraborty discloses the scene classification apparatus of video according to claim 1 or 4, wherein the video are compressed data (video source may be either compressed or decompressed video data, col. 6 lines 45-46). However, Chakraborty silent in regards to the motion intensity is detected by using a value of a motion vector of a predictive coding image existing in each shot.

However, Park teaches motion intensity is detected by using a value of a motion vector of a predictive coding image existing in each shot (column 16 line 20-35 and fig. 14).

Therefore, it would have been obvious for one of ordinary skill in the art at the time of the invention to combine the teachings of Chakraborty (modified by Toklu) with Parks' teaching of motion intensity detected by motion vectors to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have

composite information attribute, thereby being suitable for efficient data search of multimedia (column 1 line 30-36).

(2) Regarding **claim 15**, Chakraborty (modified by Toklu) as a whole teaches everything as claimed above, see claims 1 or 4. In addition, Chakraborty discloses the scene classification apparatus of video according to claim 1 or 4, wherein the video are compressed data (video source may be either compressed or decompressed video data, col. 6 lines 45-46). However, Chakraborty silent in regards to the motion intensity is detected by using a value of a motion vector of a predictive coding image existing in each shot.

However, Park teaches motion intensity is detected by using a value of a motion vector of a predictive coding image existing in each shot (column 16 line 20-35 and fig. 14).

Therefore, it would have been obvious for one of ordinary skill in the art at the time of the invention to combine the teachings of Chakraborty (modified by Toklu) with Parks' teaching of motion intensity detected by motion vectors to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have composite information attribute, thereby being suitable for efficient data search of multimedia (column 1 line 30-36).

Regarding **claim 17** Chakraborty (modified by Toklu and Park) as a whole teaches everything as claimed above, see claim9. In addition, Chakraborty discloses

the scene classification apparatus of video according to claim 9, wherein the video are compressed data video source may be either compressed or decompressed video data, col. 6 lines 45-46). Chakraborty is silent in regards to the histogram of motion direction is detected by using a value of a motion vector of a predictive coding image existing in each shot.

However, Park teaches the histogram of motion direction is detected by using a value of a motion vector of a predictive coding image existing in each shot (Park, column 16 line 63 to column 17 line 10, column 22 line 31-49, column 18 line 29-31, fig. 9 and fig. 14).

Therefore, it would have been obvious for one of ordinary skill in the art at the time of the invention to combine the teachings of Chakraborty (modified) with Parks' teaching of a histogram of motion direction is detected by using a value of a motion vector of a predictive coding image existing in each shot to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have composite information attribute, thereby being suitable for efficient data search of multimedia (column 1 line 30-36).

Regarding **claim 18**, Chakraborty (modified by Toklu and Park) as a whole teaches everything as claimed above, see claim 1 or 4. In addition, Chakraborty discloses the scene classification apparatus of video according to claims 1 or 4, wherein the video are uncompressed data (video source may be either compressed or decompressed video data, col. 6 lines 45-46). However, Chakraborty is silent in

regards to the motion intensity (is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots.

However, Park teaches the motion intensity is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots (Park, column 11, line 66 to column 12 line 7 and column 24 line 55-60, and column 18 line 29-31).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the teachings of Chakraborty (modified by Toklu) with the Parks' teaching of the motion intensity (is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots, to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have composite information attribute, thereby being suitable for efficient data search of multimedia (column 1 line 30-36).

(2) Regarding **claim 18**, Chakraborty (modified by Toklu) as a whole teaches everything as claimed above, see claims 1 or 4. In addition, Chakraborty discloses the scene classification apparatus of video according to claims 1 or 4, wherein the video are uncompressed data (video source may be either compressed or decompressed video data, col. 6 lines 45-46). However, Chakraborty is silent in regards to the motion

intensity (is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots.

However, Park teaches the motion intensity is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots (Park, column 11, line 66 to column 12 line 7 and column 24 line 55-60, and column 18 line 29-31).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the teachings of Chakraborty (modified by Toklu) with the Parks' teaching of the motion intensity (is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots, to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have composite information attribute, thereby being suitable for efficient data search of multimedia (column 1 line 30-36).

Regarding **claim 19**, Chakraborty (modified by Toklu and Park) as a whole teaches everything as claimed above, see claims 1 or 4. In addition, Chakraborty discloses the scene classification apparatus of video according to claims 1 or 4, wherein the video are uncompressed data (video source may be either compressed or decompressed video data, col. 6 line 45-46). However, Chakraborty is silent in regards to the spatial distribution of motion is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots.

However, Park teaches the spatial distribution of motion is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots (Park, column 23 line 20-30. Further Park discloses the motion direction is computed from the motion vector values, column 16 line 62-65 and column 18 line 29-31).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the teachings of Chakraborty (modified by Toklu) with Parks' teaching of spatial distribution of motion is detected by using a value of a motion vector representing a change in motion, to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have composite information attribute, thereby being suitable for efficient data search of multimedia (column 1 line 30-36).

(2) Regarding **claim 19**, Chakraborty (modified by Toklu) as a whole teaches everything as claimed above, see claims 1 or 4. In addition, Chakraborty discloses the scene classification apparatus of video according to claims 1 or 4, wherein the video are uncompressed data (video source may be either compressed or decompressed video data, col. 6 line 45-46). However, Chakraborty is silent in regards to the spatial distribution of motion is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots.

However, Park teaches the spatial distribution of motion is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots (Park, column 23 line 20-30. Further Park discloses the motion direction is computed from the motion vector values, column 16 line 62-65 and column 18 line 29-31).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the teachings of Chakraborty (modified by Toklu) with Parks' teaching of spatial distribution of motion is detected by using a value of a motion vector representing a change in motion, to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have composite information attribute, thereby being suitable for efficient data search of multimedia (column 1 line 30-36).

Regarding **claim 20** Chakraborty (modified by Toklu and Park) as a whole teaches everything as claimed above, see claims 1 or 4. In addition Chakraborty discloses the scene classification apparatus of video according to claims 1 and 4, wherein the video are uncompressed data (video source may be either compressed or decompressed video data, col. 6 line 45-46). Chakraborty is silent in regards to the histogram of motion direction (histogram difference metric) is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots.

However, Park teaches the histogram of motion direction is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots (column 11 line 14-27, column 18 line 29-31 and fig. 1J).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the teachings of Chakraborty (modified by Toklu) with Parks' teaching of spatial distribution of motion is detected by using a value of a motion vector representing a change in motion, to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have composite information attribute, thereby being suitable for efficient data search of multimedia (column 1 line 30-36).

(2) Regarding **claim 20** Chakraborty (modified by Toklu) as a whole teaches everything as claimed above, see claims 1 or 4. In addition Chakraborty discloses the scene classification apparatus of video according to claims 1 and 4, wherein the video are uncompressed data (video source may be either compressed or decompressed video data, col. 6 line 45-46). Chakraborty is silent in regards to the histogram of motion direction (histogram difference metric) is detected by using a value of a motion vector representing a change in motion predicted from a compared result of frames composing the respective shots.

However, Park teaches the histogram of motion direction is detected by using a value of a motion vector representing a change in motion predicted from a compared



result of frames composing the respective shots (column 11 line 14-27, column 18 line 29-31 and fig. 1J).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the teachings of Chakraborty (modified by Toklu) with Parks' teaching of spatial distribution of motion is detected by using a value of a motion vector representing a change in motion, to increase the speed and efficiency of data search, it has been researched and developed new search techniques which include the widely-known character-based search technique and have composite information attribute, thereby being suitable for efficient data search of multimedia (column 1 line 30-36).

6. Claims 4-6,9-14, and 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chakraborty et al., US-7, 110,454 in view of Toklu et al., US-6,549,643.

Regarding **claim 4**, Chakraborty discloses a scene classification apparatus (fig. 1) of video for segmenting video into shots (col. 5, line 1) for classifying a sequence of shots into a slow scene, where the slow scene includes a plurality of continuous shot and is thus a larger unit than a shot, comprising: an extractor for extracting from the respective shots a shot (validation module col. 7 lines 54-55) similar to a current target shot (candidate and non-candidate scene change locations (frames) col. 7 lines 36-38 and fig. 1:19) from shots after a shot before the target shot (compares neighboring keyframes col. 7 line 55) only by a predetermined interval (predetermined threshold col. 14 line 59); and a slow (gradual) scene detector (interframe variance difference col. 7

line 48-50) for classifying the target shot (candidate and non-candidate scene change locations (frames) col. 7 lines 39-38) into a slow scene (gradual) of the shot similar to the current target shot based on motion intensity (Chakraborty discloses where the output of each of the scene change detection processes are potential shots/scene change location based on the respective metrics (steps 211,212, and 213), both abrupt and gradual. The next steps in the preferred scene change detection process involve identifying and validating the scene changes based on various conditions. For instance, referring to FIG.2b, in the preferred embodiment, abrupt scene changes are identified where candidate scene change location output from the shot detection processes of the interframe and histogram difference metrics are in agreement (step 214). In particular, abrupt scene changes are identified by verifying that the conditions regarding both the interframe difference metric and the difference are satisfied. It is to be appreciated that by integrally utilizing the scene change candidates output from such shot detection processes, false alarms in identifying scene changes that may occur due to small motion where the interframe difference is high (and thus exceed the threshold in equation 11 above) will not occur since the condition for the histogram difference for the candidate must also be satisfied. (in which case, for small motion, such condition typically will not be satisfied), Column 12 line 60 to column 13 line 14 and fig. 2A-2B. Further regarding fig. 2A, element 211, outputs potential shot/scene change locations based on interframe difference, it is clear to the Examiner that Chakraborty discloses to determine if the scene is gradual or static, which reads upon the claimed limitation) of the current target shot (candidate and non-candidate scene change locations (frames)

col. 7 lines 36-38 and fig. 1:19) and the shot similar to the current target shot (key frame col. 14 lines 52-57 and fig 2B: 229), wherein each scene includes a plurality of cut points (Chakraborty teaches where a "shot" or "take" in video parlance refers to a contiguous recording or one or more video frames depicting a continuous action in time and space. Typically, transitions between shots (also referred to as "scene changes" or "cuts") are created intentionally by film directors, see col. 1 line 35-44. The examiner notes that a scene is a plurality of shots. Since a shot refers to a continuous recording of one or more video frames depicting a continuous action in time and space where the transitions between shots are called cuts and a scene is a plurality of shots, clearly, a scene is fully capable of including a plurality of shots containing multiple transitions (cuts) between the shots).

Chakraborty does not explicitly teach a shot segmentation device to segment the video into respective shots based on a cut points. However, Toklu teaches a shot segmentation device to segmentation device to segment the video into respective shots based on cut points (video segmentation module 12, column 5 line 38-57, col. 6 line 54-60, and fig 1 element 12).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Toklu with Chakraborty to generate a content based visual summary of video and facilitate digital video browsing and indexing, column 3 line 40-43).

Regarding **claim 5**, Chakraborty (modified by Toklu) as a whole teaches everything as claimed above, see claim 1. In addition Chakraborty discloses the scene classification apparatus of video according to claim 4, wherein the slow (gradual) scene detector (interframe variance difference metric computation col. 7 line 48-50 and fig. 1: 17) classifies the target shot (candidate and non-candidate scene change locations (frames) col. 7 lines 36-38 and fig. 1:19) into the slow scene (gradual scene) of the shot similar to the current target shot (candidate and non-candidate scene changes locations (frames), col. 7 line 36-38 and fig. 1:19) when the motion intensity (interframe difference col. 14 lines 30-32) of the shot similar to the current target shot (col. 7 line 36-38 and fig. 1:19) is stronger than the motion intensity (interframe difference col. 14 lines 30-32) of the current target shot (candidate and non-candidate scene change locations (frames) col. 5 line 20-24).

Regarding **claim 6**, Chakraborty (modified by Toklu) as a whole teaches everything as claimed above, see claim 1. In addition Chakraborty further discloses comprising a first highlight (gradual) scene detector (shot list database col. 8 line 8-11 fig. 1:21) for classifying a scene composed of a plurality of shots continued just before (neighboring key frames col. 7 line 55-59) the slow (gradual) scene into a first highlight (gradual) scene.

Regarding **claim 9**, Chakraborty discloses a scene classification apparatus (fig. 1) of video for segmenting video into shots (col. 5, line 1) and classifying a sequence of shots into a scene in which a camera operation has been performed, where the scene in which the camera operation has performed includes a plurality of continuous shots

and is thus a larger unit than a shot, comprising: a detector for detecting a histogram relating to motion directions of the respective shots (histogram difference metric col. 8 line 51-56 and col. 9 line 4-5); and a detector for detecting the scene in which the camera operation has been performed based on the histogram of motion direction (Chakraborty teaches where camera moves: these shots include the classical camera movements i.e. zoom, tilt, pan etc. In a preferred embodiment, the histogram difference metric given by equation (3) above is also analyzed to detect scene changes (step 206). It is also to be understood that other conventional metrics may be used for this metric such as the so called  $\chi^2$  static given by  $\chi^2$  times.

times..function..function..function..function. ##EQU00016## It is known, however, that while this statistic is more sensitive to interframe difference across a camera break, it also enhances the differences arising out of small object or camera motion, Column 11 line 35-50. Therefore, it is clear to the Examiner that Chakraborty discloses where the histogram metrics can detect camera motions, which reads upon the claimed limitation), wherein each scene includes a plurality of cut points (Chakraborty teaches where a "shot" or "take" in video parlance refers to a contiguous recording or one or more video frames depicting a continuous action in time and space. Typically, transitions between shots (also referred to as "scene changes" or "cuts") are created intentionally by film directors, see col. 1 line 35-44. The examiner notes that a scene is a plurality of shots. Since a shot refers to a continuous recording of one or more video frames depicting a continuous action in time and space where the transitions between shots are called cuts

and a scene is a plurality of shots, clearly, a scene is fully capable of including a plurality of shots containing multiple transitions (cuts) between the shots).

Chakraborty does not explicitly teach a shot segmentation device to segment the video into respective shots based on cut points. However, Toklu teaches a shot segmentation device to segmentation device to segment the video into respective shots based on cut points (video segmentation module 12, column 5 line 38-57, col. 6 line 54-60, and fig 1 element 12).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Toklu with Chakraborty to generate a content based visual summary of video and facilitate digital video browsing and indexing, column 3 line 40-43).

Regarding **claim 10**, the combination of Chakraborty as a whole teaches everything as claimed above, see claim 1. In addition Chakraborty discloses the scene classification apparatus of video according to claim 9, further comprising a zooming scene detector (interframe variance difference metric col. 4 lines 15-17) for, when the histogram of motion direction (histogram difference metric col. 8 lines 54-57) is uniform (col. 8 lines 62-63, i.e. "normal" intensity distribution) and a number of elements of respective bins is larger than a reference number of elements (each bin corresponding to an intensity range col. 8 line 53), classifying its shot into a zooming scene (gradual scene).

Regarding **claim 11**, Chakraborty (modified by Toklu) as a whole teaches everything as claimed above, see claim 1. In addition Chakraborty discloses the scene

classification apparatus of video according to claim 9, further including: detector for detecting spatial distribution (variance difference furthermore, the variance difference detects the difference within a frame where spatial distribution takes place) of motion of each shot; and a panning scene detector (interframe and histogram difference metric col. 7 lines 46-48) for detecting whether the respective shots are a panning scene (abrupt scene) based on the histogram of motion direction (histogram difference metric, the histogram as well as the interframe difference metric are processed to validate candidate scene changes as abrupt col. 7 lines 45-48 and fig. 2A: 202-203) and the spatial distribution of motion (variance difference).

Regarding **claim 12**, Chakraborty (modified by Toklu) as a whole teaches everything as claimed above, see claim 1. In addition Chakraborty discloses the scene classification apparatus of video according to claim 11, wherein when the histogram of motion (histogram difference metric) direction is concentrated in one direction and the spatial distribution (variance difference furthermore, the variance difference detects the difference within a frame where spatial distribution takes place) of motion is uniform (typically assumed not to change from frame to frame col. 12 lines 33-34), the panning scene detector (interframe and histogram difference metric col. 7 lines 46-48) classifies shot into the panning (abrupt) scene.

Regarding **claim 13**, Chakraborty discloses a scene classification apparatus of video for classifying a sequence of shots into a commercial scene, where the commercial scene includes a plurality of continuous shots and is thus larger unit than a shot, comprising: a detector for detecting a shot density DS (histogram difference

metric, a histogram is a graphical display of tabulated frequencies) of the video; and a commercial scene detector (Chakraborty discloses video are playing an increasingly import role in education and commerce, Column 1 line 15-18. Further, when the approximate maximum duration is known, since the frames/sec is always known, the maximum frame duration for the scene change is readily ascertainable. If any of the windows have a duration that exceeds this threshold, it may be assumed that the window in question is not likely to be a gradual scene change. In such as case, further examination becomes necessary. The possibilities are that either the window represents just motion or a combination of scene change and motion. In the preferred embodiment, if any window has a duration that exceeds the predefined threshold, it is assumed that the window represents motion, and consequently all points in such window are turned "off" (step 224). All the remaining windows are then identified as candidates for gradual scene change, column 14 line 20-35. Chakraborty teaches a predefined shot duration (column 13 line 15 to 35); which is equivalent to the shot density. Therefore, since Chakraborty discloses videos in education and commerce, and based on the predefined window threshold, the scene is either gradual or abrupt, it is clear to the examiner that Chakraborty is fully capable of detecting a commercial scene based on the shot density, which reads upon the claimed limitation), wherein each scene includes a plurality of cut points (Chakraborty teaches where a "shot" or "take" in video parlance refers to a contiguous recording or one or more video frames depicting a continuous action in time and space. Typically, transitions between shots (also referred to as "scene changes" or "cuts") are created intentionally by film directors, see col. 1 line 35-44. The examiner



notes that a scene is a plurality of shots. Since a shot refers to a continuous recording of one or more video frames depicting a continuous action in time and space, and a scene is a plurality of shots, clearly, a scene is fully capable of including a plurality of shots containing multiple transitions (cuts) between the shots) for detecting the commercial scene (abrupt scene) by comparing a shot density (minimum predefined shot duration col. 13 lines 18-35) detected during a predetermined interval with a predetermined reference shot density (column 14 line 21-27).

Regarding **claim 23**, see the rejection and analysis made for claim 1, except this the corresponding method claim for the apparatus of claim 1. Thus, the rejection and analysis applied for claim 1 also applies.

7. Claims 14 is rejected under 35 U.S.C. 103(a) as being unpatentable over Chakraborty et al., US-7, 110,454 in view of Toklu et al., US-6,549,643 and in view of Yilmaz et al., Shot Detection Using Principal Coordinate System.

Regarding **claim 14**, Chakraborty discloses a scene classification apparatus of video for a sequence of shots into a commercial scene, where the commercial scene includes a plurality of continuous shots and is thus a larger unit than a shot, comprising: a detector for detecting a number of shot boundaries (threshold levels, col. 5 lines 22-23, furthermore, histograms are the most common method used to detect shot boundaries) of the video; and a commercial scene detector (interframe and histogram difference metric, col. 7 lines 46-48) for detecting the commercial scene (abrupt scene Chakraborty further discloses video in education and commerce; a video in commerce

would be a commercial scene) by comparing a number of shot boundaries (threshold level col. 5 line 22-23) detected during a predetermined interval with a predetermined reference number (column 14 line 21-27), wherein each scene includes a plurality of points (Chakraborty teaches where a "shot" or "take" in video parlance refers to a contiguous recording or one or more video frames depicting a continuous action in time and space. Typically, transitions between shots (also referred to as "scene changes" or "cuts") are created intentionally by film directors, see col. 1 line 35-44. The examiner notes that a scene is a plurality of shots. Since a shot refers to a continuous recording of one or more video frames depicting a continuous action in time and space where the transitions between shots are called cuts, and a scene is a plurality of shots, clearly, a scene is fully capable of including a plurality of shots containing multiple transitions (cuts) between the shots).

Chakraborty does not explicitly teach a shot segmentation device to segment the video into respective shots based on cut points; and classifying the scene as the commercial scene in response to the comparing indicating that the number of shot boundaries detected during the predetermined interval is greater than the predetermined reference number.

However, Toklu teaches a shot segmentation device to segmentation device to segment the video into respective shots based on shot points (video segmentation module 12, column 5 line 38-57, col. 6 line 54-60 and fig 1 element 12).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Toklu with Chakraborty to generate a

content based visual summary of video and facilitate digital video browsing and indexing, column 3 line 40-43).

Chakraborty modified by Toklu is silent in regards to classifying the scene as the commercial scene in response to the comparing indicating that the number of shot boundaries detected during the predetermined interval is greater than the predetermined reference number.

However, Yilmaz teaches classifying the scene as the commercial scene in response to the comparing indicating that the number of shot boundaries detected during the predetermined interval is greater than the predetermined reference number (Yilmaz teaches to cluster news video into news and advertisements based, based on the shots boundaries detected by principle coordinate approach, we used the minimum eigenvalued eigenvector,  $v_3$ . To define a shot if it is anchor news or advertisement, we calculated the mean  $v_3$ 's in a shot and if its below a threshold, it is labeled as anchor news; otherwise it is labeled as advertisement, 4.4 Clustering Video Stream into News and Commercials. Further disclosed by Yilmaz is that shot boundaries are defined by thresholding the rotation changes for the whole video stream, 3.2 Algorithm. Therefore it is clear to the Examiner that Yilmaz teaches to determine a commercial based on the shot boundary thresholds. Since Chakraborty teaches determining a scene change (shot boundary) by comparing each of the computed metrics for the successive frames to threshold levels, and Yilmaz teaches to define a shot if it is anchor news or advertisement by the calculated  $v_3$  and a threshold, Chakraborty now (modified by Yilmaz) teaches where a commercial is determined by a shot boundary threshold, which

reads upon the claimed limitation).

Thus it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Yilmaz with Chakraborty (modified by Toklu) for improving efficiency of shot detection.

8. Claims 16 is rejected under 35 U.S.C. 103(a) as being unpatentable over Chakraborty et al., US-7, 110,454 in view of Toklu et al., US-6,549,643 and in view of Yilmaz et al., Shot Detection Using Principal Coordinate System.

Regarding **claim 16**, Chakraborty (modified by Toklu) as a whole teaches everything as claimed above, see claim 11. In addition, Chakraborty teaches wherein the video are compressed data (video source may be either compressed or decompressed video data, col. 6 lines 45-46), and the spatial distribution (variance difference, referring to within the frame, furthermore, MPEG has spatio temporal locator capabilities) of motion is detected by using a value of a motion vector of a predictive coding image existing in each shot (MPEG, col. 6 lines 51-60, furthermore, MPEG is a predictive image coding technique that incorporates tabulating motion vector values).

10. Claims 7-8 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chakraborty (US Patent 7,110,454) in view of Toklu et al., US-6,549,643 and in further view of Blanchard US Patent 6347114).

Regarding **claim 7**, Chakraborty fails to teach a detector for detecting the intensity of audio signals accompanied by the video. Blanchard teaches a detector for detecting intensity of an audio signal (audio levels col. 3 lines 37-51) accompanied by the video (col. 2 lines 27-29) into shot. Blanchard also teaches detector for classifying a

scene composed of a plurality of shots continued before and after a shot with the audio signal intensity stronger than the predetermined intensity (col. 2 lines 17-22) into a second highlight scene (gradual scene).

Taking the combined teaching of Chakraborty (modified by Toklu) and Blanchard as a whole, it would have been obvious to one of ordinary skill in the art at the time that the invention was made to incorporate detecting the intensity of audio signals accompanied by the video as claimed for the benefit of detecting scene changes that may generally be identified and distinguished from mere shots changes where the audio level will generally remain the same.

Regarding **Claim 8**, the combination of Chakraborty (modified by Toklu) and Blanchard as whole further teaches everything claimed as applied above; see claims 7. In addition Chakraborty teaches a commercial scene detector (interframe and histogram difference metric col. 7 lines 46-48, Chakraborty) for classifying the respective shots into a commercial scene (abrupt scene), wherein a scene classified into a scene other than the first highlight scene (gradual), the second highlight scene (gradual scene) and the commercial scene (abrupt scene) is classified into the highlight scene (gradual).

12. Claim 21 is rejected under 35 U.S.C 103(a) as being unpatentable over Nakamura et al., US-2001/0051516 and in view of Pan et al., US-2002/0080162 in view of Gonsalves et al., US-6,392,710 and further in view of Chakraborty et al., US-7,110,454.

Regarding **claim 21**, Nakamura teaches a scene classification apparatus of video for segmenting video into shots based on cut points and classifying each scene

composed of one or more continuous shots based on a content of the scene comprising: a detector for detecting a highlight scene (In such a case that a plurality of highlight scenes are detected by the analyzing unit 22, [0208] and fig. 2); extracting and combining means for extracting and combining a plurality of highlight scenes (In such a case that a plurality of highlight scenes are detected by the analyzing unit 22 from a program during a CM broadcasting time range, and the present CM broadcast is commenced, the reproducing management unit 27 reproduces a plurality of detected highlight scenes in a time sequential manner by equally increasing a reproducing speed, [0208] and fig. 2. Nakamura discloses the reproducing management unit 27 reproduces a plurality of detected highlight scenes and the highlight scenes are stored in a highlight scene index storage unit, (fig. 2, element 25), it is clear to the examiner that in order to reproduce the highlight scenes stored in the storage unit, by the reproducing management unit, the highlight scenes are retrieved and combined, thus reading upon the claimed limitation). Nakamura is silent in regards to inserting means for inserting a video transition effect into a combined portion of the respective highlight scenes, the inserting means including a dynamic/static scene detector to detect whether a highlight scene is a dynamic scene with much motion or a static scene with little motion wherein the inserting means makes a type of the video transition effect to be inserted different according to whether the highlight scenes to be combined are they dynamic scene or the static scene.

However, Pan teaches inserting means for inserting a video transition effect into a combined portion of the respective highlight scenes, the inserting means including a

dynamic/static scene detector to detect whether a highlight scene is a dynamic scene with much motion or a static scene with little motion (Pan teaches where the pattern of a slow motion replay in sports program including very fast movement of objects (persons, ball, and etc.), generally referred to as action shots at block 10. Following the action shots at block 10 there may be other shots or video content at block 12 prior to the slow motion replay segment in block 14. A special effect, or edit at block 16, is almost always present between the normal shots in block 12 and 16 and the slow motion replay segment in block 18. After the slow motion replay in block 18, another edit effect in block 20, is usually present before resuming normal play. A more detailed structure of the slow motion replay 14 of FIG. 1 is shown in FIG. 2. Typically the procedure of the slow motion replay includes six components, namely, edit effects in 20, still fields 22, slow motion replay 24, normal replay 26, still fields 28, and edit effect out 30, [0028-0029]. The edit effects in 20 and edit effects out 30, mark the starting and end points of the procedure of the slow motion replay 14, and typically are gradual transitions, such as fade in/out, cross/additive-dissolve, and wipes. Frequently, the logo of the television station will be shown during the edit effects in 20 and edit effects out 30. Other techniques may likewise be used, col. Therefore, it is clear to the Examiner that Pan discloses an inserting means to insert a transition effect into the action shots, and determines if the action shot is a slow replay shot or normal speed replay, which reads upon the claimed limitation.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Pan with Nakamura for providing a

generic system for video analysis which reliably detects semantically significant events in a video, [0006].

Nakamura (modified by Pan) is silent in regards to wherein the inserting means makes a type of video transition effect to be inserted different according to whether the highlight scenes to be combined are the dynamic scene or the static scene.

However, Gonsalves teaches allowing the video editor to insert a video transition effect on a field/frame-by-field/frame basis in order to improve accuracy of the effect (Gonsalves, special effect, col. 3 line 11-14 line 24, between two frames col. 4 line 65-67, col. 5 lines 50-52, and fig. 3b: 320a-320b).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Gonsalves with Nakamura (modified by Chakraborty) to improve accuracy of the effect.

Nakamura (modified by Pan and Gonsalves) is silent in regards to wherein each scene includes a plurality of cut points.

However, Chakraborty teaches where a "shot" or "take" in video parlance refers to a contiguous recording or one or more video frames depicting a continuous action in time and space. Typically, transitions between shots (also referred to as "scene changes" or "cuts") are created intentionally by film directors, see col. 1 line 35-44. The examiner notes that a scene is a plurality of shots. Since a shot refers to a continuous recording of one or more video frames depicting a continuous action in time and space where the transitions between shots are called cuts and a scene is a plurality of shots,



clearly, a scene is fully capable of including a plurality of shots containing multiple transitions (cuts) between the shots.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Chakraborty with Nakamura (modified by Pan and Gonsalves) for providing an efficient method for video browsing and content based retrieval.

Claim 22 is rejected under 35 U.S.C 103(a) as being unpatentable over Nakamura et al., US-2001/0051516 and in view of Pan et al., US-6,931,595 in view of Gonsalves et al., US-6,392,710 and further in view of Gotoh et al., US-5,801,765.

Regarding **claim 22**, Nakamura (modified by Pan and Gonsalves) as a whole teaches everything as claimed above, see claim 21. Nakamura is silent in regards to the scene classification apparatus of video according to claim 21, wherein when the highlight scene is the dynamic scene, the video transition effect with small change in an image mixing ratio is inserted therein, and when the highlight scene is the static scene, the video transition effect with large change in the image mixing ratio is inserted therein.

However, Gotoh discloses where specifically, the scene-change is classified into two types depending on how a video changes: one in which a scene changes momentarily; and one in which a scene changes gradually. Those generally referred to as the scene-change is the former, i.e., a scene appeared in a moment of pressing a record start button (see Fig. 11(a)). The latter are those given special effects, such as effect and fade, when editing a video (see Fig. 11 (b)). Hereinafter, the former and the latter are referred to as "momentary scene-change" and "gradual scene-change"

respectively. In a gradual scene-change, it takes much time that a scene change to another. In the Fig. 11 (b), pictures H to K comprise a gradual scene-change, column 2 line 17-29 and fig. 11(a) and 11(b). Therefore, it is clear to the examiner that Gotoh discloses a special effect that has a momentary change for a dynamic scene and a special effect that takes much time to change for gradual scene, which reads upon the claimed limitation.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate the teachings of Gotoh with Nakamura (modified by Pan and Gonsavles) for providing improved image quality.

### ***Conclusion***

9. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

***Contact Information***

Any inquiry concerning this communication or earlier communications from the examiner should be directed to JESSICA PRINCE whose telephone number is (571)270-1821. The examiner can normally be reached on 7:30-5:00 EST Monday-Friday, Alt Friday off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Marsha D. Banks-Harold can be reached on (571) 272-7905. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Marsha D. Banks-Harold/  
Supervisory Patent Examiner, Art Unit 2482  
/JESSICA PRINCE/  
Examiner, Art Unit 2482